# ‖‖‖‖‖‖‖‖‖‖‖‖‖‖

PCT10

**RAW SEQUENCE LISTING**                    DATE: 09/12/2002
**PATENT APPLICATION:  US/10/030,829**       TIME: 15:54:19

Input Set : **A:\100303829SeqList.txt**
Output Set: **N:\CRF4\09122002\J030829.raw**          ENTERED

```
 1 <110> APPLICANT: Beclin, Christophe
 2         Elmayan, Taline
 3         Vaucheret, Herve
 5 <120> TITLE OF INVENTION: NOVEL SGS3 PLANT GENE AND USES THEREOF
 8 <130> FILE REFERENCE: A34920-PCT-USA 072667.0179
10 <140> CURRENT APPLICATION NUMBER: 10/030,829
11 <141> CURRENT FILING DATE: 2002-01-11
13 <150> PRIOR APPLICATION NUMBER: PCT/FR/00/02052
14 <151> PRIOR FILING DATE: 2000-01-26
16 <150> PRIOR APPLICATION NUMBER: FR 99/09,417
17 <151> PRIOR FILING DATE: 1999-07-16
19 <160> NUMBER OF SEQ ID NOS: 5
21 <170> SOFTWARE: FastSEQ for Windows Version 4.0
23 <210> SEQ ID NO: 1
24 <211> LENGTH: 3275
25 <212> TYPE: DNA
26 <213> ORGANISM: Arabidopsis thaliana
28 <220> FEATURE:
29 <221> NAME/KEY: primer_bind
30 <222> LOCATION: (693)...(715)
31 <223> OTHER INFORMATION: p356AD'
33 <221> NAME/KEY: primer_bind
34 <222> LOCATION: (2926)...(2952)
35 <223> OTHER INFORMATION: P356Y'
```

W-->  37 <400> 1

```
38 gacaaacaaa caaaaattaa gcaagtcatg ttcgtagcaa taaattaata gtgggaacaa 60
39 ttaagttaag cgaaaaagga aaaaaaaagg tacaaaaatg aaaacaaaat caaactgaat 120
40 gaaaatttgg agtccagaat cggaaaaacg aggccgtttt agagcttaat aagcttcctc 180
41 atttgtctct tcttcgtcag tttattttct tcctccagag tcctgactca ctactctcac 240
42 tctcggcgc tttaaactta cgttctccgt cgtttactct gtaagttttc tgccttagag 300
43 cctccgatcg cctcaccgca tgcattctgt gctcgatttc tcttttcctt cgctggaaaa 360
44 attgccctaa tgttctcgat ttcgaaggtt tttgtgctat gggttacttt tttccctata 420
45 ttttatagtt cttaggtaac gatacctgcg tcttactgtt tttgttcatt ttgttgtgct 480
46 ttcaccgttt agtcgctgat cggagtattt gactgtgaaa aatccttcgt tttttggttt 540
47 ttgtttcata taaatcggat tgatctacct tttgtgcttt gatgtttgtt ttttgagcct 600
48 atgcgttgtt ggcttgttat aacttcacgt tcatgtgtgg attttgagat tttggtagtg 660
49 actgtgggtt tctttggtgg ctataggttg taaaaatgag ttctagggct ggtccaatgt 720
50 ctaagggaaa gaacgttcag ggtggttata ggcctgaggt tgaacagttg gttcaaggtt 780
51 tggcagggac gagactggct tcttcacaag atgatggagg agagtgggag gtcatttcca 840
52 agaagaacaa gaacaaacca ggaaacactt ctggaaaaac ttgggtttct cagaattcga 900
53 atcctcctag agcttggggt ggtcagcagc aagggagagg tagcaacgta tctgggagag 960
54 gaaacaatgt atccgggaga ggtaacggca atggtcgggg cattcaagct aacatatctg 1020
55 gtcggggacg agcgttgagc agaaagtatg ataacaactt tgtggcaccc ccacctgtat 1080
```

RAW SEQUENCE LISTING
PATENT APPLICATION·   US/10/030,829

DATE: 09/12/2002
TIME: 15:54:19

Input Set : A:\100303829SeqList.txt
Output Set: N:\CRF4\09122002\J030829.raw

```
56 ctcgccctcc tttggaagga ggatggaatt ggcaggcaag aggaggttct gctcagcaca 1140
57 cagctgtgca ggagtttcct gacgtggagg atgatgtgga taatgcttct gaggaagaga 1200
58 atgattccga tgctttggat gattctgatg acgaccttgc aagtgatgat tatgactcgg 1260
59 atgtgagtca aaagagccat ggatcacgaa agcagaataa gtggttcaaa aagttctttg 1320
60 gcagcttgga tagcttgtcg atcgagcaga taaatgaacc acagaggcag tggcattgtc 1380
61 cagcttgtca gaacggacct ggtgccatcg attggtataa cctgcaccct ctactagctc 1440
62 atgcgaggac aaaaggagct aggcgagtta agctccatag agaattggct gaagttttag 1500
63 aaaaggatct acagatgaga ggcgcatctg tcattccttg tggtgagatt tatgggcagt 1560
64 ggaagggttt gggtgaggat gaaaaggatt atgaaattgt ctggcctcca atggtcatca 1620
65 tcatgaatac tagactggat aaggacgata acgataaggt ggaattcttc tgtctttttac 1680
66 ttctttaatt tttctcttgc attctactga tcttagaatg ttacattgta gtggctcggc 1740
67 atgggcaacc aagagctgct ggaatacttc gacaagtatg aggctcttag agcacgccat 1800
68 tcctatggtc cacagggcca tcgtgggatg agtgttctga tgtttgagag cagtgccact 1860
69 ggctatttgg aggccgaacg cctccaccgg gagttagctg agatgggggtt agatagaatt 1920
70 gcctgggggtc agaagcgcag tatgttttct ggaggtgttc gccaactgta tggcttcctt 1980
71 gcaacgaagc aagatctgga catattcaat caacactctc aaggttctct cccccaaaga 2040
72 aatttgatat atgctttttag ttttgtcatt ggaatttaaa gttttgttgg tccgtgttaa 2100
73 tgcatctgtt atgtatatat ctatgattca ttaggcaaaa caaggctgaa attcgagttg 2160
74 aaatcatacc aagagatggt tgtaaaggag ctgaggcaga tctctgagga caatcagcag 2220
75 ctgaactact ttaagaacaa gctctcaaaa cagaacaagc acgccaaggt gcttgaggaa 2280
76 tctctggaaa ttatgagcga gaagctgcgt agaactgcag aggataatcg gatcgtgaga 2340
77 cagagaacta agatgcagca tgaacagaac agggaagagg tatgattttt cctagaaaat 2400
78 cacaaacttg acattttgta ttacctactg attcacattt ttgattatat tgtccaacaa 2460
79 aaaacctgtg gtggtttgaa gatggatgca cacgacaggt ttttcatgga ttcaatcaaa 2520
80 cagatccatg aaagaagaga cgcaaaggag gagaatttcg agatgttgca gcagcaggaa 2580
81 cgtgccaagg ttgttggcca gcagcagcag aacattaatc cctctagcaa tgacgattgc 2640
82 cgaaagaggt atatgtacta actaacataa tccctctggc gtttttgttt ttcaaaccta 2700
83 agagtaactg aattattccg gttttgattc tttcgcagag ctgaggaagt gtcaagcttc 2760
84 atcgagtttc aagagaaaga gatggaggag tttgtggaag agaggagat gctgataaaa 2820
85 gatcaagaga agaagatgga agacatgaag aagaggcatc acgaggagat atttgatctg 2880
86 gagaaagaat ttgatgaggc tttggaacag ctcatgtaca agcatggcct tcacaatgaa 2940
87 gatgattgag acaaaagtct ggtacacaag acaagactaa gtttctttgt tttgcttttg 3000
88 gtatgtcgga aagtaggaga tctgagagac tccatttaaa tactaggaca aatctaagga 3060
89 gattatagat tattatcctc caatttttag tagacggatc taaggaagca ttaagttctt 3120
90 gtgactaaaa ccaagtttcc ttagtatttt gttttttttt ggtaaaattt catatgaaag 3180
91 ttagacatat taccaaacgt cagagtgaat cacagaatgg caaatcaaaa tcatgttttt 3240
92 agaattttat atctacaaaa tatatgggta caaat                            3275
94 <210> SEQ ID NO: 2
95 <211> LENGTH: 1878
96 <212> TYPE: DNA
97 <213> ORGANISM: Arabidopsis thaliana
99 <220> FEATURE:
100 <221> NAME/KEY: CDS
101 <222> LOCATION: (1)...(1878)
103 <400> SEQUENCE: 2
104 atgagttcta gggctggtcc aatgtctaag gaaaagaacg ttcagggtgg ttataggcct 60
105 gaggttgaac agttggttca aggtttggca gggacgagac tggcttcttc acaagatgat 120
106 ggaggagagt gggaggtcat ttccaagaag aacaagaaca aaccaggaaa cacttctgga 180
107 aaaacttggg tttctcagaa ttcgaatcct cctagagctt ggggtggtca gcagcaaggg 240
```

RAW SEQUENCE LISTING                    DATE: 09/12/2002
PATENT APPLICATION:   US/10/030,829     TIME: 15:54:19

Input Set : A:\100303829SeqList.txt
Output Set: N:\CRF4\09122002\J030829.raw

```
108 agaggtagca acgtatctgg gagaggaaac aatgtatccg ggagaggtaa cggcaatggt 300
109 cggggcattc aagctaacat atctggtcgg ggacgagcgt tgagcagaaa gtatgataac 360
110 aactttgtgg cacccccacc tgtatctcgc cctcctttgg aaggaggatg gaattggcag 420
111 gcaagaggag gttctgctca gcacacagct gtgcaggagt ttcctgacgt ggaggatgat 480
112 gtggataatg cttctgagga agagaatgat tccgatgctt tggatgattc tgatgacgac 540
113 cttgcaagtg atgattatga ctcggatgtg agtcaaaaga gccatggatc acgaaagcag 600
114 aataagtggt tcaaaaagtt ctttggcagc ttggatagct tgtcgatcga gcagataaat 660
115 gaaccacaga ggcagtggca ttgtccagct tgtcagaacg gacctggtgc catcgattgg 720
116 tataacctgc accctctact agctcatgcg aggacaaaag gagctaggcg agttaagctc 780
117 catagagaat tggctgaagt tttagaaaag gatctacaga tgagaggcgc atctgtcatt 840
118 ccttgtggtg agatttatgg gcagtggaag ggtttgggtg aggatgaaaa ggattatgaa 900
119 attgtctggc ctccaatggt catcatcatg aatactagac tggataagga cgataacgat 960
120 aagtggctcg gcatggggaa ccaagagctg ctggaatact tcgacaagta tgaggctctt 1020
121 agagcacgcc attcctatgg tccacagggc catcgtggga tgagtgttct gatgtttgag 1080
122 agcagtgcca ctggctattt ggaggccgaa cgcctccacc gggagttagc tgagatggggg 1140
123 ttagatagaa ttgcctgggg tcagaagcgc agtatgtttt ctggaggtgt tcgccaactg 1200
124 tatggcttcc ttgcaacgaa gcaagatctg gacatattca atcaacactc tcaaggcaaa 1260
125 acaaggctga aattcgagtt gaaatcatac caagagatgg ttgtaaagga gctgaggcag 1320
126 atctctgagg acaatcagca gctgaactac tttaagaaca agctctcaaa acagaacaag 1380
127 cacgccaagg tgcttgagga atctctggaa attatgagcg agaagctgcg tagaactgca 1440
128 gaggataatc ggatcgtgag acagagaact aagatgcagc atgaacagaa cagggaagag 1500
129 atggatgcac acgacaggtt tttcatggat tcaatcaaac agatccatga aagaagagac 1560
130 gcaaaggagg agaatttcga gatgttgcag cagcaggaac gtgccaaggt tgttggccag 1620
131 cagcagcaga acattaatcc ctctagcaat gacgattgcc gaaagagagc tgaggaagtg 1680
132 tcaagcttca tcgagtttca agagaaagag atggaggagt ttgtggaaga gagggagatg 1740
133 ctgataaaag atcaagagaa gaagatggaa gacatgaaga agaggcatca cgaggagata 1800
134 tttgatctgg agaaagaatt tgatgaggct ttggaacagc tcatgtacaa gcatggcctt 1860
135 cacaatgaag atgattga                                                  1878
137 <210> SEQ ID NO: 3
138 <211> LENGTH: 625
139 <212> TYPE: PRT
140 <213> ORGANISM: Arabidopsis thaliana
142 <400> SEQUENCE: 3
143 Met Ser Ser Arg Ala Gly Pro Met Ser Lys Glu Lys Asn Val Gln Gly
144  1               5                   10                  15
145 Gly Tyr Arg Pro Glu Val Glu Gln Leu Val Gln Gly Leu Ala Gly Thr
146                 20                  25                  30
147 Arg Leu Ala Ser Ser Gln Asp Asp Gly Gly Glu Trp Glu Val Ile Ser
148             35                  40                  45
149 Lys Lys Asn Lys Asn Lys Pro Gly Asn Thr Ser Gly Lys Thr Trp Val
150     50                  55                  60
151 Ser Gln Asn Ser Asn Pro Pro Arg Ala Trp Gly Gly Gln Gln Gln Gly
152 65                  70                  75                  80
153 Arg Gly Ser Asn Val Ser Gly Arg Gly Asn Asn Val Ser Gly Arg Gly
154                 85                  90                  95
155 Asn Gly Asn Gly Arg Gly Ile Gln Ala Asn Ile Ser Gly Arg Gly Arg
156                 100                 105                 110
157 Ala Leu Ser Arg Lys Tyr Asp Asn Asn Phe Val Ala Pro Pro Pro Val
158             115                 120                 125
```

RAW SEQUENCE LISTING                    DATE: 09/12/2002
PATENT APPLICATION:   US/10/030,829     TIME: 15:54:19

Input Set : A:\100303829SeqList.txt
Output Set: N:\CRF4\09122002\J030829.raw

```
159 Ser Arg Pro Pro Leu Glu Gly Gly Trp Asn Trp Gln Ala Arg Gly Gly
160     130                 135                 140
161 Ser Ala Gln His Thr Ala Val Gln Glu Phe Pro Asp Val Glu Asp Asp
162 145                 150                 155                 160
163 Val Asp Asn Ala Ser Glu Glu Glu Asn Asp Ser Asp Ala Leu Asp Asp
164                 165                 170                 175
165 Ser Asp Asp Asp Leu Ala Ser Asp Asp Tyr Asp Ser Asp Val Ser Gln
166                 180                 185                 190
167 Lys Ser His Gly Ser Arg Lys Gln Asn Lys Trp Phe Lys Lys Phe Phe
168         195                 200                 205
169 Gly Ser Leu Asp Ser Leu Ser Ile Glu Gln Ile Asn Glu Pro Gln Arg
170     210                 215                 220
171 Gln Trp His Cys Pro Ala Cys Gln Asn Gly Pro Gly Ala Ile Asp Trp
172 225                 230                 235                 240
173 Tyr Asn Leu His Pro Leu Leu Ala His Ala Arg Thr Lys Gly Ala Arg
174                 245                 250                 255
175 Arg Val Lys Leu His Arg Glu Leu Ala Glu Val Leu Glu Lys Asp Leu
176                 260                 265                 270
177 Gln Met Arg Gly Ala Ser Val Ile Pro Cys Gly Glu Ile Tyr Gly Gln
178     275                 280                 285
179 Trp Lys Gly Leu Gly Glu Asp Glu Lys Asp Tyr Glu Ile Val Trp Pro
180     290                 295                 300
181 Pro Met Val Ile Ile Met Asn Thr Arg Leu Asp Lys Asp Asp Asn Asp
182 305                 310                 315                 320
183 Lys Trp Leu Gly Met Gly Asn Gln Glu Leu Leu Glu Tyr Phe Asp Lys
184                 325                 330                 335
185 Tyr Glu Ala Leu Arg Ala Arg His Ser Tyr Gly Pro Gln Gly His Arg
186                 340                 345                 350
187 Gly Met Ser Val Leu Met Phe Glu Ser Ser Ala Thr Gly Tyr Leu Glu
188     355                 360                 365
189 Ala Glu Arg Leu His Arg Glu Leu Ala Glu Met Gly Leu Asp Arg Ile
190     370                 375                 380
191 Ala Trp Gly Gln Lys Arg Ser Met Phe Ser Gly Gly Val Arg Gln Leu
192 385                 390                 395                 400
193 Tyr Gly Phe Leu Ala Thr Lys Gln Asp Leu Asp Ile Phe Asn Gln His
194                 405                 410                 415
195 Ser Gln Gly Lys Thr Arg Leu Lys Phe Glu Leu Lys Ser Tyr Gln Glu
196                 420                 425                 430
197 Met Val Val Lys Glu Leu Arg Gln Ile Ser Glu Asp Asn Gln Gln Leu
198     435                 440                 445
199 Asn Tyr Phe Lys Asn Lys Leu Ser Lys Gln Asn Lys His Ala Lys Val
200     450                 455                 460
201 Leu Glu Glu Ser Leu Glu Ile Met Ser Glu Lys Leu Arg Arg Thr Ala
202 465                 470                 475                 480
203 Glu Asp Asn Arg Ile Val Arg Gln Arg Thr Lys Met Gln His Glu Gln
204                 485                 490                 495
205 Asn Arg Glu Glu Met Asp Ala His Asp Arg Phe Phe Met Asp Ser Ile
206             500             505             510
207 Lys Gln Ile His Glu Arg Arg Asp Ala Lys Glu Glu Asn Phe Glu Met
```

```
208           515                   520                   525
209 Leu Gln Gln Gln Glu Arg Ala Lys Val Val Gly Gln Gln Gln Gln Asn
210       530                   535                   540
211 Ile Asn Pro Ser Ser Asn Asp Asp Cys Arg Lys Arg Ala Glu Glu Val
212 545                   550                   555                   560
213 Ser Ser Phe Ile Glu Phe Gln Glu Lys Glu Met Glu Glu Phe Val Glu
214               565                   570                   575
215 Glu Arg Glu Met Leu Ile Lys Asp Gln Glu Lys Lys Met Glu Asp Met
216           580                   585                   590
217 Lys Lys Arg His His Glu Glu Ile Phe Asp Leu Glu Lys Glu Phe Asp
218           595                   600                   605
219 Glu Ala Leu Glu Gln Leu Met Tyr Lys His Gly Leu His Asn Glu Asp
220       610                   615                   620
221 Asp
222 625
225 <210> SEQ ID NO: 4
226 <211> LENGTH: 23
227 <212> TYPE: DNA
228 <213> ORGANISM: Artificial Sequence
230 <220> FEATURE:
231 <223> OTHER INFORMATION: Oligonucleotide p356AD'
233 <400> SEQUENCE: 4
234 aaaatgagtt ctaggctgg tcc                                          23
236 <210> SEQ ID NO: 5
237 <211> LENGTH: 27
238 <212> TYPE: DNA
239 <213> ORGANISM: Artificial Sequence
241 <220> FEATURE:
242 <223> OTHER INFORMATION: Oligonucleotide p356Y'
244 <400> SEQUENCE: 5
245 gtctcaatca tcttcattgt gaaggcc                                     27
```

**VERIFICATION SUMMARY**                         DATE: 09/12/2002
PATENT APPLICATION:    US/10/030,829             TIME: 15:54:20

Input Set : **A:\100303829SeqList.txt**
Output Set: **N:\CRF4\09122002\J030829.raw**

L:37 M:258 W: Mandatory Feature missing, <220> not found for SEQ ID#:1